

KNOWLEDGE AND NEUROSCIENCE

Comments on Churchland's "What Happens to Reliabilism When It Is Liberated from the Propositional Attitudes?"

Joe Cruz
Williams College

Let me begin with the standard apology and expression of regret for not being able to comment on all of the intriguing and illuminating themes in Professor Churchland's paper. I should at least note, though, my enthusiasm for his suggestive discussion of the complexity of all concepts, for his detailed portrayal of the resources of neural network models, and for his attempt to deflate our Cartesian pretensions by focusing on the commonality between human and infrahuman cognition.

I restrict my developed remarks to two different topics: First, I have some thoughts about Churchland's central thesis with respect to the nature of epistemology. Second, I want to say something about the pursuit of epistemology if we follow Churchland's lead.

I

Churchland draws to our attention a range of capacities that enable our manifestly spectacular success in navigating the world. These include, naming just a few, those capacities that give us the perceptual skill to recognize faces, or locomotive gaits, or

spatial relations. According to Churchland, the philosopher cannot ship these mechanisms off to neurophysiology for illumination, cheerful in the confidence that they have nothing to do with epistemology. On the contrary, these mechanisms are alleged to be of just the right sort to play a role in naturalistic epistemological theories. The thought is that, when we take seriously the advice that epistemology needs to include on its agenda the processes causally implicated in epistemic achievements, Churchland's mechanisms seem to need to be on the table.

Now, one concern I have here is that the plausibility of this conclusion comes from tacitly reading the processes as yielding *beliefs* such as *that is John's face*, or *that is a human face*, or *that is a forward moving feline conspecific* (a reading that Churchland no doubt thinks is 'romantic nonsense'). But tacitly introducing attitudinal output to the description of the process will not do for Churchland's purpose. He needs for these mechanisms to seem plausibly epistemological on a non-propositional reading.

My suspicion, though, is that we will accept a capacity as a candidate for filling the crucial causal role in epistemology precisely to the degree that we can successfully reconcile its output with the propositional attitude approach. There is no doubt that an attitudinal characterization of cognitive science is contentious, but one of the reasons cognitive science has been attractive to naturalized epistemologists is that it seems to be amenable to a treatment of its paradigmatic processes in attitudinal terms.

This picture is prominent in Goldman's treatment of these topics. Goldman's strategy has been to pursue a refinement of the propositional attitudes in light of cognitive scientific research. In *Epistemology and Cognition*, for instance, Goldman defends a liaison between psychology and epistemology on the basis of psychology's promise of illuminating the causes of *belief*.

Of course, this is just what Churchland is complaining about, and it is hardly a reply to his work for us to say that our prejudices in favor of belief are deep. I am not convinced, however, that they are *merely* prejudices. The traditional, propositional attitude epistemology has some resources for mounting a reasoned defense.

What seems to be at issue here is what we take as the criterion for treating a process or mechanism as occupying a crucial role in understanding specifically epistemic achievements. Putting it this way would apparently, though not surprisingly, require us to recognize epistemic achievements *in advance* of casting about for the appropriate processes. How should we carry out this identificatory project? One natural suggestion, and the answer given by generations of epistemologists, is to focus on intuitive mundane cases of belief formation that constitute our everyday commitments about the world. Thus we have an initial constraint on what sorts of conversations we may have as epistemologists. That closes the door to some processes, namely the ones that do not issue forth in belief or belief-like attitudes.

This is not to say that we couldn't revise our conception of the horizon of epistemology, or that we couldn't develop and defend a broader framework within which the mundane cases constitute a subset. This lesson, too, has been prominent in Goldman's work. Again, his version has always been along the lines of refining our conception of the kinds of processes that are analyzable at the level of the propositional attitudes. In the same spirit, Churchland is trying to widen our vision, but he needs to tell us why an interest in mechanisms characterized at a neurological level is demanded by our goals *as epistemologists*. It surely will not do, for example, to point out that the integrity and reliability of the neurological processes is a prerequisite for the kinds of perceptual knowledge that we have. The integrity and reliability — in some sense of these terms — of our endocrine system is essential to our perceptual knowledge. So, for that matter, is the integrity and reliability of physical systems in general. In the bizarre world where we become periodically 'unglued' at the atomic level, perceptual knowledge as we understand it is impossible. Naturalistically inclined epistemologists do not, however, need to attend to the endocrine system, or to physics. Our interest in causal mechanisms cannot derive *solely* from the fact that the reliability of the processes is a necessary precondition for knowledge.

Let us step back to the point where our inclination was to tacitly or explicitly redescribe the output of perceptual processes in propositional attitude form. Suppose we could resist this redescription, and could confidently defend the non-attitudinal version of

the processes as appropriate elements of epistemology. We would still have to question the rationale for closing ourselves off from an explanatory framework that resonates with our traditional methodology in epistemology. Churchland claims that attitudinal redescription is not an option in many cases, because there is no redescription to be had of these skills for infrahumans. He cautions that, in these contexts, “the familiar homocentric, linguaformal, propositional attitude conception...functions only poorly, or only metaphorically, if it functions at all” (p. 3).

The argument for this claim seems to be that, for many activities that require knowledge (broadly construed), it is either not credible that propositional attitudes are formed or that it is indeterminate how many or which ones were formed. I take it that this is the force of the lathe and drill press case, the driving case, and the basketball player case.

Churchland says, “In the course of a silent hour of industry at my workshop...my perceptual adventures are considerable.... But I may not have fixed a single propositional attitude in the course of the entire hour” (p. 4). Here, he seems to be pointing out that there is no occasion of a conscious thought in sentence or quasi-sentence form and there is certainly not any talking. Fair enough. To say the least, though, it is extremely contentious whether conscious sentence-form thoughts exhaust the realm of the propositional attitudes.

This is one of the nagging elements of Churchland's presentation. In several places he points out the absence of language as impugning the propositional attitude approach. I am not inclined to treat speech or its silent internal conscious analog as a crucial source of our confidence in propositional attitudes. One might instead sign onto their explanatory credentials by claiming, say, that the inferentially governed rule-like forms (algorithms, perhaps) are descriptions of what the underlying architecture realizes. One could be an instrumentalist or realist about this story, but either way, it seems like the propositional attitude approach does not hang on the presence of language.

By way of the indeterminacy point, Churchland points out that any attempt at treating his examples within the propositional attitude framework would face embarrassing questions. For instance, we would have to offer an answer to "how *many* discrete P-A's does our perceptual machinery fix per second?" And our answers would have to be grounded "in a detailed understanding of the brain's microstructure..." (p. 4).

But the demand that the answer be grounded in the brain's microstructure is clearly a partisan one that we are entitled to reject. For any given cognitive process, cognitive psychology needs to countenance only as many propositional attitudes as are required for satisfying explanations of the behavioral evidence. Put so baldly, this may sound irremediably tepid. But I take it that this is nearly enough the received methodology of cognitive psychology. We start with the data acquired in carefully controlled experiments, and we postulate mechanisms that would account for that data in

information processing terms. I would not have thought that it should be a source of anxiety that cognitive psychologists might be unable, for instance, to distinguish between a hypothesis where one propositional attitude is the output of a perceptual system versus the hypothesis where a thousand identical propositional attitudes are serially produced in a kind of streaming pulse of output over the span of 500ms. Unless distinguishing these two hypotheses becomes crucial within the norms of psychological explanations, this is simply not an issue that cognitive psychology needs to address.

At any rate, this seems to me what many cognitive psychologists and epistemologists have in mind when they invoke the propositional attitudes. There may be many things illicit about this perspective, but it does not seem that any of Churchland's remarks here can be taken as addressing it head-on.

II

I turn next to a different kind of concern. If we take Churchland's advice on epistemology — and especially if we are persuaded that it is not just knowledge-how but knowledge in general that must yield to neuroscientific insight — the normative enterprise of epistemology would be dramatically reconfigured. I submit that epistemology becomes so dauntingly esoteric that it becomes unclear why we would want to do it anymore.

Churchland's conception of success for cognitive representations is when the 'distance' relations between attractors in the hyper-dimensional space instantiated by weights and activation values mirror the 'distance' relations in the target real-world domain. The mirroring here is an isomorphism of similarity relations between the activation space and the environment, and the accuracy of the representations is a function of the *degree* of mirroring.

Naturally, we are now very far from allowing folk intuitions to govern the outcome of epistemology. But worse, we are very far from allowing folk intuitions to govern the aspirations, questions, or intelligibility of the results of epistemology, since the details of the success of the mapping will be available only to the specialist. It won't suffice for the neuroscientifically less informed epistemologist to note the success of cases of knowledge-how. After all, a person could be representing a domain in a wildly inaccurate way that nonetheless allows them to succeed systematically in their projects. Think of a decent chess player who is convinced that the late game of chess is governed by all sorts of bizarre rules, like the rule that rooks and pawns can arbitrarily switch places on the 60th move (suppose that their early experiences with chess were with a really buggy chess computer). But imagine that this player's games rarely go beyond 60 moves, so it seems to you that she is pretty good at chess. The person represents chess wildly inaccurately, but they do okay, so only a neuroscientist could find out.

To see how deep this feature of the account is, notice that the encrypted conceptual spaces are not bound by anything in our high-level descriptions of epistemic achievement. There is no reason to believe that the contours of the sculpted conceptual space will respect the boundaries of any classifications we recognize. Churchland has cheated a bit in presenting the cognitive capacities in a tidy form. He talks about recognizing faces, for instance, presumably because it is a good bet that this capacity is subsumed by a domain specific mechanism. Many of the rest of our know-how capacities will not admit to this characterization.

The basketball star whose representational space encodes the knowledge of how to score will not likely have attractors that correspond to intelligible things like *zone defense*, or *excellent shot blocker in the paint*. This means that we can't say that "Jones knows how to score against a zone defense or knows what to do when there is an excellent shot blocker in the paint" without severely distorting the reality of the situation with our "paleolithic" folk categories. Even if it makes sense to talk about basketball skill in isolation from a myriad of other non-basketball skills, we have no vocabulary for characterizing the nature of the domain that a player's how-to knowledge successfully represents. Indeed, it is not clear that we can say anything about a player's know-how, except to gesture at one of Churchland's histograms and utter, "she knows how to do the thing that is represented by that, but we have no words for what that skills is." Perhaps

we can add that “we suspect it has something to do with good basketball, but only very detailed neuroscientific investigation can find that out.”

Take another example. Consider the mundane but putatively epistemological questions, *does Carlos know how to recognize good tomatoes in the produce section?* Remember that even seemingly pedantic answers are ruled out in Churchland’s view. We cannot adequately answer, if Churchland is right, that Carlos believes that good tomatoes are firm and shiny, that he has an intact perceptual system, and that the mechanism that accesses his memory of good making tomato properties is reliable. Instead, we must investigate the activation space that Carlos exploits for his selection of tomatoes. The attractors in that space, encrypted and morphed through his long experience shopping in grocery stores, may not, need not, and probably will not correspond to simple things like firmness and shininess. Unless we think that Carlos has a dedicated domain specific tomato recognizer (and we don’t), there will be massive interaction between vectors representing experiences of produce in general, experiences of anxiety before dinner parties, experiences of produce at particular markets, etc. The conceptual space created by these complicated influences will not make sense if we try to force a middle-size object level story on them.

So, again, we can’t — without distortion — say what Carlos knows when he successfully picks out good tomatoes.

Thus, epistemology in Churchland's style is exhausted by mapping the activation space and marveling without comment on how it allows successful navigation through reality. Epistemology is neuroscience and the rest of us should go home. No wonder philosophers have preferred what they might take to be a more revealing shorthand, namely the expression of epistemic achievements in terms of propositional attitudes.